Proceedings of ASA Section on Statistics in Epidemiology
ORISE-07-OEWH-0234

Age-Based Methods to Explore Time-Related Variables in Occupational Epidemiology Studies

Janice P. Watkins[1], Edward L. Frome[2], Donna L. Cragle[1]
[1]Center for Epidemiologic Research, ORAU
[2]Oak Ridge National Laboratory

Abstract

Although age is recognized as the strongest predictor of mortality in chronic disease epidemiology, a calendar-based approach is often employed when evaluating time-related variables. An age-based analysis file, created by determining the value of each time-dependent variable for each age that a cohort member is followed, provides a clear definition of age at exposure and allows development of diverse analytic models. To demonstrate methods, the relationship between cancer mortality and external radiation was analyzed with Poisson regression for 14,095 Oak Ridge National Laboratory workers. Based on previous analysis of this cohort, a model with ten-year lagged cumulative radiation doses partitioned by receipt before (dose-young) or after (dose-old) age 45 was examined. Dose-response estimates were similar to calendar-year-based results with elevated risk for dose-old, but not when film badge readings were weekly before 1957. Complementary results showed increasing risk with older hire ages and earlier birth cohorts, since workers hired after age 45 were born before 1915, and dose-young and dose-old were distributed differently by birth cohorts. Risks were generally higher for smoking-related than non-smoking-related cancers. It was difficult to single out specific variables associated with elevated cancer mortality because of: (1) birth cohort differences in hire age and mortality experience completeness, and (2) time-period differences in working conditions, dose potential, and exposure assessment. This research demonstrated the utility and versatility of the age-based approach.

Key Words: occupational epidemiology, time-related variables, radiation, age at exposure

Introduction

Observational studies are an important source of information on potential adverse health effects from exposure to radiation and other hazardous materials, and great interest centers on the evaluation and incorporation of time-related variables. Prior to beginning analysis, researchers must make decisions that may have a profound effect on study outcomes, most notably characterizing annual radiation dose for use as a predictor variable in statistical models. A distinctive feature of this study was a new tool for Poisson regression that uses age rather than calendar year as the time scale. An electronic notebook (Geist and Nachtigal 2000) accessible on the Internet contains abundant details of this study, along with script files and data files. http://www.csm.ornl.gov/~frome/abaf/

Radiation cohort studies commonly utilize Poisson regression (Frome et al., 1973, Breslow and Day, 1987, Frome et al., 1997). Certain studies have supported the premise that radiation doses received by nuclear workers at older ages are associated with elevated risk for cancer mortality, while other research has found no such relationship. Cardis et al. (1995) found no difference in risk by age at exposure in a combined cohort International Agency for Research on Cancer (IARC) study. In contrast, Kneale and Stewart (1995), using combined U.S. cohorts, reported increased cancer risk after age 70 from doses received after age 50. Richardson and Wing (1998) examined several methods for relating doses to exposure age and concluded that partitioning dose at a specific age was preferable. Their Oak Ridge National Laboratory (ORNL) cohort was included in both the IARC and Kneale and Stewart studies and provided a source file for the current research. This source file contained one record for each of 14,095 ORNL cohort members and included annual external radiation doses from 1943 through 1985, available because ORNL began monitoring workers at risk for exposure from the beginning of plant operation (Watkins et al. 1993; Watkins et al. 1997).

Researchers have used many time-related variables as potential risk factors or effect modifiers in cohort studies. Age at risk is essential since background mortality rates rise with age. Other factors of interest have included active worker status, age at first exposure, age at hire, age at last exposure, annual cumulative dose, annual dose, birth cohort, calendar period, dose in age-at-exposure windows, dose in time-since-exposure windows, length of employment, length of follow-up, time since first exposure, and time since last exposure. Many of these variables are highly interrelated or even aliased, complicating assessment of their relationships with each other and dose. Researchers must choose among these variables since it is not possible mathematically to estimate parameters using models where all are included simultaneously.

Methods

Age Based Analysis Files (ABAF)

The ABAF approach evaluates key time-dependent variables annually. For each cohort member a separate record is created for each age from hire age to age when follow-up ceased, associating a dose with each age. For example, a worker with birth date 10/1/1920 would have age 40 dose equal to one-fourth the 1960 dose plus three-fourths the

1961 dose. Adjustments are needed for ages when employed only part of a year, such as first and last years of employment. Details for creating ABAFs are available (Watkins et al. 2004). The ORNL ABAF contained 435,061 records for 425,400 person-years. Dose for each age was cumulative dose with no lag applied, although alternative dose definitions could have been used.

Analytic Data Structures (ADS)

Analysis-specific ADSs, which are multidimensional tables for grouped data analysis, can be created from an ABAF. ADSs with doses lagged 10 years were created to examine these research questions: differences between age-based and calendar-year-based results; issues when partitioning dose into dose-young and dose-old; differences between multiplicative relative risk (MRR) and excess relative risk (ERR) estimates and between estimates using a baseline model and stratifying on factors; effect modification from hire age groups and birth cohorts; dose-response before and after 1957; cancers related and not related to smoking. Results for additional research questions appear in the electronic notebook and the full report (Watkins et al. 2004). ADSs with additional time-related variables of interest could have been created in the same manner.

The specific combination of factor levels (including dose category) for each record in the ABAF determined its assignment to a specific cell in the ADS. Total person-years, cancer deaths, mean age, and mean doses were calculated for each cell. Because of the high degree of stratification from factor-level combinations, using mean dose per cell approximated a continuous measure of dose. To compare with the calendar-year based approach, for age at risk younger than 55 (age 45 with doses lagged 10 years) dose was assigned to dose-young and dose-old was zero. For age at risk 55 or older, dose-young remained constant and dose-old was cumulative dose minus dose-young. An example of the assignment of the lag 10 dose to total dose, dose-young, and dose-old categories appears in Table 1.

Data Analysis

Analyses were conducted using the AMFIT module of EPICURE (Preston et al. 1993), S-PLUS (S-Plus 6 for Windows, 2003) and R (R Development Core Team, 2004). AMFIT was used to examine differences in dose-response estimates caused by specifying models as either MRR or ERR and by either selecting particular main effects and interactions as a baseline model or stratifying on the same set of selected factors. All dose-response estimates are MRR increases or decreases per 10 mSv for all cancer deaths with cumulative doses lagged 10 years, except as otherwise specified. Definitions of all analysis variables appear in Table 2. The basic forms of the models are listed below. In models for certain research questions HG, XP, or AG were used along with, or instead of, factors listed in the baseline models below.

Table 1. Example dose assignment for a worker hired at age 36, terminating at age 51, receiving 10 mSv per year occupational dose, and dying at age 57.

Age	cum. dose no lag	cum. dose 10-yr lag	dose-young[a] 10-yr lag	dose-old[b] 10-yr lag
36	10	0	0	0
37	20	0	0	0
38	30	0	0	0
39	40	0	0	0
40	50	0	0	0
41	60	0	0	0
42	70	0	0	0
43	80	0	0	0
44	90	0	0	0
45	100	0	0	0
46	110	10	10	0
47	120	20	20	0
48	130	30	30	0
49	140	40	40	0
50	150	50	50	0
51	160	60	60	0
52	160	70	70	0
53	160	80	80	0
54	160	90	90	0
55	160	100	90	10
56	160	110	90	20
57	160	120	90	30

[a]Dose-young: Cumulative dose received before age 45;
[b]Dose-old: Cumulative dose received at age 45 or older

MRR with baseline model:

Rate = exp(LA+GE+RC+PA+CH+WK+FA+IX +LA*WK+GE*RC+GE*CH+PA*CH **+dose**)

ERR with baseline model:

Rate = exp(LA+GE+RC+PA+CH+WK+FA+IX +LA*WK+GE*RC+GE*CH+PA*CH) **(1+dose)**

MRR with stratification:

Rate = b_a exp(**dose**)

ERR with stratification:

Rate = b_a**(1+dose)** , where the b_a are separate multiplicative parameters for each level of each factor in the baseline model with LA being replaced by the factor AG.

Table 2. Definitions of variables used in analysis.

LA	Natural logarithm of (age/55)
AG	5-Year Age Groups beginning with < 25 and ending with 90+
GE	Gender: level 1 male, level 2 female
RC	Race: level 1 white, level 2 non-white
PA	Pay code: level 1 weekly, level 2 hourly, level 3 monthly (surrogate for socio-economic status)
CH	Birth cohort: level 1 1915-24, level 2 1905-14, level 3 < 1905, level 4 ≥1925
WK	Active worker: level 1 not active, level 2 active (2-year lag)
FA	Ever a multiple facility worker: level 1 no, level 2 yes
IX	Internal exposure: level 1 not monitored (low exposure potential), level 2 monitored (higher exposure potential), level 3 not eligible (before 1951 when internal monitoring began) (10-year lag)
DT	Total cumulative dose (10-year lag)
DY	Dose-young; dose received before age 45 with a 10-year lag (i.e. cum dose received before age 55)
DO	Dose-old; dose received at age 45 or older with a 10-year lag
HG	Hire-age: level 1 < 30 years, level 2 30-44, level 3 45 or older
XP	Exposure period: level 1 pre 1957, level 2 1957 + (10-year lag)

Results

Distributions of Person-Years and Cancer Deaths

Fig. 1 shows the similarity of person-years distributions for total dose and dose-young in contrast to dose-old, for which 91% of all person-years were assigned a zero dose. Fig. 2 reveals a somewhat closer correspondence between dose-young and dose-old in dose-group assignment of cancers, although a larger number of cancers were assigned to a dose-old of zero. Comparing the highest dose group between dose-young and dose-old, the difference in person-years is substantial while the difference in cancer deaths is relatively slight.

Age-Based Compared To Calendar-Year-Based Results

Dose-response estimates for age-based results using the baseline model and the MRR were 0.0560 for dose-old (s.e. 0.0139) and -0.0115 for dose-young (s.e. 0.0109), and age-based results with stratification were nearly identical. The MRR baseline model calendar-year estimates (Wing and Richardson 1997) were dose-old of 0.0498 (s.e. 0.0148), which was substantially the same as the age based result,

Figure 1: Person-years by dose-young, dose-old, and total cumulative dose

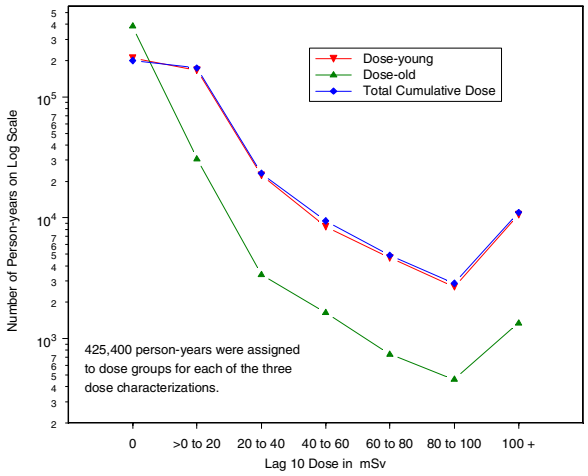
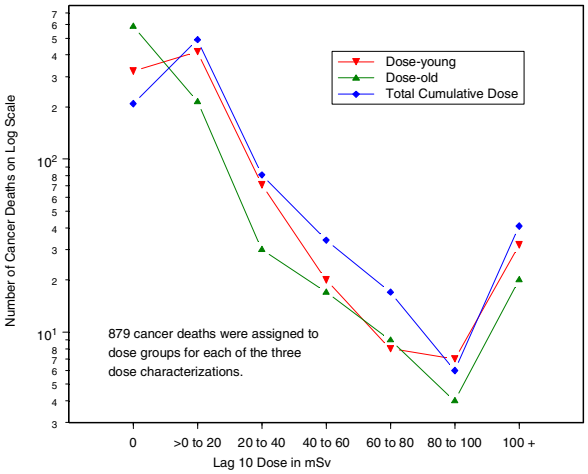


Figure 2: Cancer deaths by dose-young, dose-old, and total cumulative dose



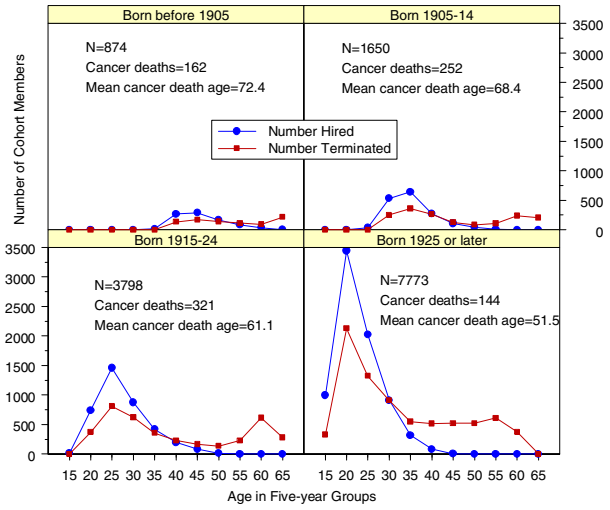
and dose-young estimate of -0.0069 (s.e. 0.0105), which was within a standard error of the age-based estimate. The ERR age-based estimates and standard errors were somewhat larger but the same in pattern and likely to have resulted in similar interpretations.

Hire Age and Birth Cohort

There were extended opportunities for workers hired before age 30 to receive occupational radiation doses before age 45. In contrast, workers hired after age 45 did not receive dose before age 45. Fig. 3 shows the strong relationship between hire age and birth cohort since most cohort members were hired before age 30 and born after 1915, while individuals hired after age 45 were born before 1915.

Although more than 82% of the workers were born after 1915, information about the cancer mortality experience of these younger workers, was much less complete than for older workers. From oldest to youngest birth cohort, the percents who died of cancer were 18.5%, 15.3%, 8.5%, and 1.9%, and the latest ages for which cancer deaths could be identified were 100, 85, 75, and 65 years, respectively. Hire age and birth cohort were also linked to dose-young and dose-old. Dose-response estimates in Table 3 show an apparent trend for cancer mortality risk with increasing hire age. However, only the 1905-14 birth cohort had substantially elevated risk.

Figure 3: Hire age and termination age* by birth cohort



*Age at end of study used if still employed in 1990.

Table 3. Dose-response estimates[a] for hire age groups.

Name	Estimate	Std.Err	TestStat	P value
Total cumulative dose (lag 10)				
Hire age group				
HG_1*DT (< 30)	−0.0097	0.0125	−0.781	0.435
HG_2*DT (30-44)	0.0208	0.0080	2.590	0.010
HG_3*DT (45+)	0.0852	0.0369	2.309	0.021
Birth cohort				
CH_4*DT (1925+)	−0.0135	0.0213	−0.633	> 0.500
CH_1*DT (1915-24)	−0.0066	0.0108	−0.617	> 0.500
CH_2*DT (1905-14)	0.0315	0.0080	3.955	< 0.001
CH_3*DT (pre 1905)	0.0227	0.0303	0.750	0.453

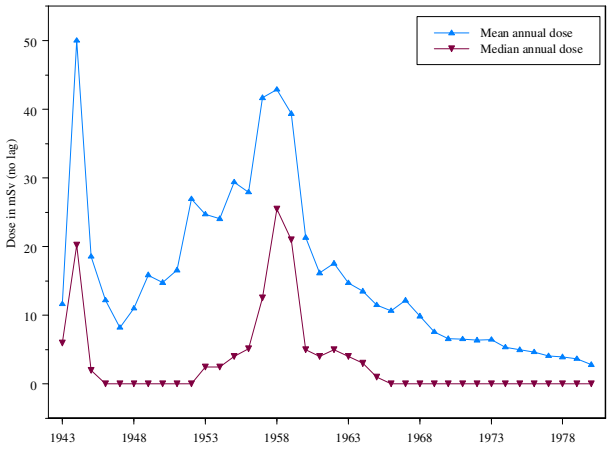
[a]Increase or decrease in ln(RR) per 10 mSv.

Exposure Period

Exposure period was dichotomized at each worker's age in 1957, the year when ORNL began reading film badges quarterly rather than weekly, allowing more time for weekly doses to accumulate and reach the observable threshold for detection. A 10 year lag in exposure period allowed evaluation of the impact of differences in dose assessment during the two time periods. Fig. 4 shows that in the transition year of 1957 and the two following years, doses were distinctly elevated above all other years except for 1944. Beginning in 1960, workers were less likely to receive substantial radiation doses, and over one-third of the cohort was hired in 1960 or later.

There was a suggestion of a positive association between total dose and cancer risk in the later time period (0.01, s.e. 0.007) but no association between dose and cancer before 1957 (-0.02, s.e. 0.08), an attenuation which could be due to dose underestimation. Partitioning dose at age 45 revealed opposite patterns of dose-young and dose-old in the two time periods.

Figure 4: Average annual external dose*



*In mSv with no lag applied.

Cancers Related To and Not Related To Smoking

Smoking-related cancers had ICD-8 codes 140-149 (mouth and oral cavity), 150 (esophagus), 157 (pancreas), 161 (larynx), 162 (lung), 188 (bladder), 189.0-189.2 (kidney, ureter). Table 4 shows that parameter estimates for smoking-related cancers were larger than non-smoking-related while standard errors were similar. In particular, the dose-old estimate for cancers not related to smoking was less than half as large as the estimate for smoking-related cancers, suggesting the possibility of a birth cohort effect.

Table 4. Dose-response estimates[a] for smoking-related and

non-smoking-related cancers.

Name	Estimate	Std.Err	TestStat	P value
Total cumulative dose (lag 10)				
DT (smoking)	0.0140	0.0089	1.583	0.113
DT (non-smoking)	0.0059	0.0106	0.555	> 0.500
Cumulative dose partitioned at age 45 (lag 10)				
DY (smoking)	−0.0216	0.0166	−1.297	0.195
DO (smoking)	0.0736	0.0206	3.581	< 0.001
DY (non-smoking)	−0.0061	0.0166	−0.366	> 0.500
DO (non-smoking)	0.0297	0.0253	1.176	0.239

[a]Increase or decrease in ln(RR) per 10 mSv.

Discussion

This study investigated the issue of time-related variables in occupational cohort studies starting with an age-based analysis file, in which dose and key time-related variables are evaluated for each year of a worker's age rather than calendar year. Analyses to demonstrate methods were based on a 14,095 member ORNL cohort with follow-up from 1943 through 1990. The ABAF approach proved to be a useful tool for providing a clear operational definition of dose received at a given age and facilitating the creation of various ADSs to address research questions including whether radiation doses received at later ages are related to increased risk of dying of cancer. This age-based approach can be used with a variety of source data files, such as those in which annual doses are adjusted for minimum detection level or those with only basic time-related variables such as birth date, current age, and dose.

Complementary results showed increasing risk with older hire ages and earlier birth cohorts since generally workers hired before age 30 were born after 1915 while those hired after age 45 were born before 1915 (Fig. 4). These patterns in parameter estimates mirrored the dose-response from partitioning total dose at age 45 since dose-young and dose-old were distributed differently for various hire ages and birth cohorts. Only the 1905-14 birth cohort had substantially elevated mortality. The dose-old estimate for smoking-related cancers was more than twice as large as the estimate for non-smoking-related cancers, suggesting the possibility of a birth cohort effect.

Because of the complex, interconnected relationships among time-related variables, it was not possible to pinpoint the most suitable set of time-related variables for analysis in this cohort. Possible risk factors for cancer mortality likely provided different levels of risk to various birth cohorts. These factors included the age at which doses were received, differences in lifestyle and environment between earlier and later birth cohorts, and differences in working conditions, exposure potential, and exposure assessment from early to later years of plant operation. Also, much higher percentages of workers from the older birth cohorts were deceased by 1990, making their cancer mortality experience more complete than for younger birth cohorts. Lower average doses for cohort members with increasing calendar time complicated relationships among time-related variables, as did the differing distribution patterns of dose-young and dose-old among the various hire age groups and birth cohorts.

Since birth cohort is an important modifier of radiation risk in the ORNL workforce, extending follow-up for all birth cohorts would facilitate more complete evaluation of effect modification by age at exposure. The study of other occupational radiation-exposed cohorts with wider birth year distributions could help to elucidate the relative importance of birth cohort and exposure age as modifiers of radiation risk. The age-based approach will facilitate continuing investigations with time-related variables in occupational studies.

Acknowledgements

This research was conducted by the Oak Ridge Institute for Science and Education with funding from the Health-Related Energy Research Branch (HERB), National Institute for Occupational Safety and Health under Contract Number 200-1999-08009. The work was done in collaboration with Mary Shubauer-Berigan and Sharon R. Silver of HERB, whose contributions are gratefully acknowledged. This work has been authored by a contractor of the U.S. Government. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this work, or to allow others to do so for U.S. Government purposes.

References

Breslow NE, Day NE. Statistical Methods in Cancer Research, Vol. II – The design and Analysis of Cohort Studies. IARC Sci Publ. No. 82, IARC, Lyon; 1987.

Cardis E, Gilbert E, Carpenter L, et al. Effects of low doses and low dose rates of external ionizing radiation: cancer mortality among nuclear industry workers in three countries. Radiat Res. 142:117-132; 1995.

Frome EL, Kutner MH, Beauchamp JJ. Regression analysis of Poisson-distributed data. J Am Stat Assoc. 68:935-940; 1973.

Frome EL, Cragle DL, Watkins JP, et al. A mortality study of employees of the nuclear industry in Oak Ridge, Tennessee. Radiat Res. 148:64-80; 1997.

Geist A, Nachtigal N. Oak Ridge National Laboratory electronic notebook project at ORNL, Oak Ridge, TN; 2000.

Kneale GW, Stewart AM. Factors affecting recognition of cancer risks of nuclear workers.Occ Envir Med.52(8):515-23;1995.

Preston D, Lubin J, Pierce D, and McConney M. Epicure User's Guide. Hirosoft International Corp., Seattle; 1993.

R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3; 2004.

Richardson D, Wing S. Methods for investigating age differences in the effects of prolonged exposures. Am J Ind Med. 33:123-130; 1998.

S-Plus 6 for Windows, Version 6.2. Seattle: Insightful Corp.; 2003.

Watkins JP, Reagan JL, Cragle DL, et al. Data collection, validation, and description of data for the Oak Ridge nuclear facilities mortality study. ORAU Technical Report 93/J-42; 1993. http://www.orau.gov/cer/reports.htm

Watkins JP, Cragle DL, Frome EL, et al. Collection, validation, and treatment of data for a mortality study of nuclear industry workers. Appl Occup Environ Hyg. 12:195-205; 1997.

Watkins JP, Frome EL, Cragle DL. Final report: Evaluating time-related variables in occupational epidemiology studies. NIOSH – HERB Contract Number 200-1999-08009; 2004.

Wing SB, Richardson D. Final report: Time-related factors in radiation-cancer dose-response. NIOSH Grant No. R03 OH03343; 1997.